Validating a habitat evaluation method for predicting avian richness

Paul R. Adamus

Abstract A new avian richness evaluation method (AREM) was developed and tested for assessing lowland wetland and riparian habitats of the Colorado Plateau. AREM rapidly scores habitats for avian richness from simple observations of habitat characteristics. AREM's predictions were compared with original field data from 76 sites on the Colorado Plateau during the breeding season. Species predictions and detections were highly indicative of the breeding avifauna in regional wetlands studied. AREM has implications for use in mitigation calculations, detection of impaired wetland quality, selection of appropriate indicator species, targeting habitat enhancements, wildlife-based classification of wetland habitats, and assisting strategies for protecting biodiversity.

Key words birds, Colorado, diversity, habitat assessment, habitat–relationship models, methods, riparian, wetlands

Planners and wildlife managers must often assign relative values to habitat patches to support objectives for project mitigation and habitat restoration, management, and planning. Detailed, standardized procedures such as those proposed by Golet (1973) and the widely-used Habitat Evaluation Procedure (HEP) of the U.S. Fish and Wildlife Service (1980) are among many methods that address this need. More general habitat classification schemes are also used that couple information on cover-class distribution with predictions of habitat suitability, using simple wildlife–habitat–relationship (WHR) models (Morrison et al. 1992). Although either of these approaches can improve the explicitness, consistency, and scientific basis for decisions, current methods for evaluating habitat have several limitations (Van Horne and Wiens 1991). These include lack of clear definition of what is being measured (which species, attributes, and at which scales), over-reliance on a few purported "indicator" species, failure to explicitly address biological diversity, and perhaps most importantly, a paucity of validation (i.e., testing of model accuracy).

I describe herein the results of testing one community-level habitat model, the avian richness evaluation method (AREM), during the breeding season. AREM was developed to improve existing habitat assessment approaches by alleviating some of their limitations. The primary objective of field testing was to determine the agreement, at 76 wetland and riparian sites, between observed avian richness and avian richness predicted using AREM's onsite habitat assessments and models. Avian richness was a focus because it is a component of biodiversity which is recognized in legislation and in policies of several resource agencies. A secondary objective of field testing was to determine simi-

Author address: Paul R. Adamus, ManTech Environmental Research Services Corporation, USEPA Environmental Research Laboratory, 200 Southwest 35th Street, Corvallis, OR 97333, USA.

Peer refereed

Colorado Plateau area for which the avian richness evaluation method (AREM) was developed and region (arrow) where AREM was tested, May–June 1993.

larity between species composition observed and species composition predicted using AREM's 56 species models that represent virtually all birds known to regularly nest in lowland wetland and riparian areas of the Colorado Plateau (map), as determined from literature and discussions with local naturalists (Kingery 1988, Dexter and Lavad 1992). Explicitly evaluating accuracy of individual species models was a minor objective. I anticipated that some species models, due to weak information, might be too conservative or liberal, but if model predictions for individual species were summed for all species, the resulting estimate of avian richness would be realistic.

Methods

Model development and characteristics

AREM is a series of WHR models, 1 for each of 165 bird species regularly inhabiting lowland riparian and wetland areas of the Colorado Plateau. The model for each species describes habitat features (indicators) that determine habitat suitability or are correlated with presence of that species. Models were developed based mainly on my experience and on interviews with local avian experts. Each WHR model uses site-specific habitat information to assign a "species habitat score" from 0 (least suitable habitat) to 1 (most suitable habitat) for a species. After visit-

ing a site, a user answers standardized questions (see box for condensed version). The user processes the responses on a PC (personal computer)-based program (described in Adamus 1993*a*) which compares the site's habitat conditions with a database that defines habitat preferences of each species. The computer program tallies species with scores above various threshold (cutoff) values in the 0–1 scale. For example, with a score cutoff of >0.75, only species for which the site is particularly suitable are included. Unless noted otherwise, results reported herein are based on the >0.75 cutoff.

AREM's output variables include number of species (avian richness), sum of all the species' scores, and sum of species' scores individually weighted by various characteristics (e.g., neotropical migrant status). These outputs help assign priorities among individual wetlands or wetland complexes. The software also allows users to review and edit models for any species.

Testing

Seventy-six wetland and riparian sites at elevations of 1,487–1,844 m (4,880–6,050 ft) were selected in a portion of the Colorado Plateau in Montrose and Delta Counties, Colorado. Wetlands in this area are relatively accessible, and some were mapped (U.S. Bureau of Reclamation 1991). Sites were selected based on accessibility rather than statistically because the study was intended to validate AREM, not to characterize the wetland population.

Questions asked for the avian richness evaluation method (AREM) which predicted avian species composition and richness in 76 wetland and riparian habitats (0.04–3 ha) of the Colorado Plateau. This list of features is complete, but descriptions are greatly condensed from the actual 6-page AREM field form in Adamus (1993a), which also includes existing literature documentation.

Landscape scale

Proximity to major river, lake: <0.9 km to river >30 m wide or lake >16 ha?

Land cover types: within 0.9 km, cover is >60% agriculture and wetland, or desert, or pinyon–juniper (*Pinus-Juniperus* spp.), or oak, or other.

Predator-facilitating land cover patterns: great (urban, by major road, or a linear patch), moderate (other road or building within 305 m), or other.

Seclusion: great (>183 m from road *and* on-foot human visits infrequent), moderate (>183 m from road *or* on-foot human visits infrequent), or other

Site scale

Surface water: >0.04 ha?

Open water: large (>8 ha and >152 m wide), small (<0.04 ha, or >0.04 ha and <1 m wide), or other?

Still water: >0.04 ha flowing at <0.3 m/sec?

Fish, amphibians, crayfish: known or expected to be present?

Water transparency: <25 cm or feedlot adjoins site?

Drawdown: normally inundated but dries out or floods from river 1 year in 5?

Bare soil-mud: large (>0.4 ha, >30 m wide, not alkali, not topographically recessed), moderate (>0.04 ha), or other?

Tree cover: >2 trees onsite? within 305 m?

Snags, large trees: >2 snags or large trees (>2 cm diameter) within 91 m?

Shrub cover: extent of willow (*Salix* spp.), greasewood (*Sarcobatus* spp.), fleshy-fruited shrubs, or tamarisk (*Tamarix* spp.); in open or dense stands; and in narrow or wide patches (size categories 0.04–0.4 ha or >0.4 ha).

Herbaceous cover: extent of robust emergents, other wet emergents, or drier emergents; in open or dense stands; and short or tall; and in narrow or wide patches (size categories 0.04–0.4 ha, or >0.4 ha).

Nesting structures: platforms, sand banks within 0.9 km?

Grazing, mowing, burning: intensive removal of cover?

Most sites were <0.2 ha (range 0.04–3 ha), lacked surface water, adjoined small roads, and were surrounded by cropland or rangeland. Most contained >0.04 ha of a combination of cattail (*Typha latifolia*) and shrubs (especially greasewood [*Sarcobatus* spp.] and willow [*Salix* spp.]). Trees were present at 38% of the sites, but 85% of the stands of trees covered <0.4 ha. All sites have been partly or completely sustained for decades by artificially high water tables caused by irrigation runoff (Rector et al. 1979, Adamus 1993b). Ten of the sites were constructed wetlands or wetlands that had been altered (generally within 5 years) to enhance suitability to waterfowl.

Bird species were identified by sight or vocalization by an observer standing or walking through sites during early morning hours and under favorable weather conditions, 29 May–20 June 1993. Local sources believed the sample year and months to be much wetter than usual. Search time was standardized as 5 minutes/10 m of wetland site length. Most (79%) sites were visited twice. At most sites the 2 visits combined exceeded 30 minutes. One observer conducted all surveys. Soras (*Porzana carolina*) and Virginia rails (*Rallus limicola*) were surveyed by playing taped vocalizations (Glahn 1974) during the initial 5 minutes of the survey. Nocturnal species were not included because the field methods could not detect these species effectively.

Categorical information on habitat features (box) was collected using AREM's standard field form at each site. To ensure comparability of predicted and observed values, the habitat characterizations (which were the basis for the predictions) were performed

only within the area of each site that was surveyed for birds, regardless of whether any birds were found within this area.

Data analysis

The suite of species models and databases contained in the AREM program developed prior to field surveys was used to convert information on the AREM field forms to predictions of avian richness and species composition at each site. Although AREM models were developed for 165 species, I only tested models for the 56 species known to breed in the region's lowland riparian and wetland sites. The data included predictions of 56 species and observations of 59 species, indexed by site. Lists of observed species and predicted species were aggregated by site and species and assigned to 1 of 4 categories. *Positive matches* were species predicted to be present that were, in fact, observed. *Null matches* were species predicted to be absent that were not detected (i.e., not observed). *Commissions* were species predicted to be present that were not detected (Type I error). *Omissions* were species predicted to be absent that were present (Type II error). These categories (number of positive matches, null matches, commissions, and omissions) were tallied by species and by site.

Accuracy of individual species models was estimated by subtracting number of sites with omission or commission errors from number of sites with null or positive matches. Variation among sites in error rates was determined by subtracting, at each site, number of species with omission or commission errors from number of species with null or positive matches. Overall accuracy was estimated using the Spearman rank correlation test if a difference existed between a ranking of sites based on observed and on predicted avian richness. Similarity of observed and predicted species composition was estimated using the simple matching coefficient (SMC, Sokal and Michener 1958). The SMC is more appropriate than other indices when null matches and positive matches are evidence of model validity (Digby and Kempton 1987). The SMC was used to compare predicted species compositions with observed species composition at each site.

$$S = m/(c + o + m) \qquad (1)$$

where S = simple matching coefficient, m = number of predicted species that were observed (both positive and null matches), c = number of predicted species that were not observed (commissions), and o = number of species observed but not predicted (omissions).

Results

Overall accuracy

Predicted richness values ($\bar{x}$ = 8.87, SE = 0.823, n = 76) did not differ from observed values ($\bar{x}$ = 10.58, SE = 0.688, n = 76) as indicated by the Wilcoxon signed ranks test (Z = 3.546, 2-tailed probability of exceeding Z = 0.0004). The Spearman rank correlation test, comparing predicted and observed richness across the 76 sites, was positive, relatively large (r^2 = 0.80), and highly significant ($P \leq 0.001$). This indicated that the distribution of richness scores, not just their means, was similar between observed and predicted data. Predicted and observed species compositions, as indicated by the SMC, also were similar. SMC values ranged from 0.74–1.00 on a possible scale of 0 (no similarity) to 1 (complete similarity) and had a median of 0.85 ($\bar{x}$ = 0.88, SE = 0.015, n = 76).

Accuracy by site and species

Of the 56 species predicted, 4 were not observed; of the 59 species observed, 7 were not predicted. The model did not predict these species because AREM contained no models for them (prior to the survey, I did not consider them to breed in the habitat and subregion with sufficient regularity to be included).

Mean scores of species–site combinations that represented positive matches ($\bar{x}$ = 0.71, SE = 0.007, n = 623) were greater than means of species-site combinations that represented commission errors ($\bar{x}$ = 0.67, SE = 0.004, n = 1,616), according to the Mann-Whitney U-test (Z = 4.25, $P \leq 0.001$, n = 2,239). Higher scores for correctly predicted species than for incorrectly predicted ones are expected if species models are generally accurate, and this is what I found.

Number of correct predictions (positive plus null matches) exceeded failures (omissions plus commissions) at all of the 76 sites and for all but 4 (7%) of the 56 species. Of the 56 species predictions made at each site, an average of 6% (3.1 species, SE = 0.5) were positive matches, 82% (46.0, SE = 1.1) were null matches, and 11% (6.2, SE = 0.5) were commission errors. The large proportion of null matches was expected, because in any region only a few species are widespread and many have restricted distributions (Preston 1948, Ricklefs and Schluter 1993). If species were predicted more liberally by using a cutoff of >0 for individual species scores, there were 19% positive matches, 41% null matches, and 36% commission errors. At this cutoff, only 1 site had experienced more AREM failures than successes. Omis-

sion errors exceeded commission errors for 21% (12 of 56) of the species at the >0 cutoff and for 32% (18 of 56) of the species at the >0.75 cutoff.

Avifaunal characteristics

Although not intended to represent a probabilistic sample of regional avifauna, these data generally demonstrated the relative frequency of occurrence of bird species in lowland wetland and riparian areas of the Colorado Plateau. Cumulatively, species found at the 76 study sites comprise virtually all of the subregion's species that breed in those habitats and >80% (59 of 73) of species that breed regularly in any habitat within the Colorado Plateau (Kingery 1988, Dexter and Lavad 1992). Although 1 site contained only a single species, most contained >9 species (13% of Colorado Plateau species) and 1 site contained 29 (40% of Colorado Plateau species). All of the 76 sites contained at least 1 neotropical migrant species, and 1 site contained 21. Neotropical migrants comprised 25–100% (median = 75%) of species observed at a site.

Discussion

Biologists sometimes disagree as to what constitutes a sufficient degree of accuracy and what parameters should judge this accuracy (e.g., Garrison 1993). For example, Rice et al. (1986:79) stated, "In general, models must be correct much more often than two-thirds of the time before one may conclude that true bird–habitat relationships are reflected by prediction," while Hurley (1986:152) suggested, "...managers are quite comfortable with accuracy levels of 75–80% for total model output..." Because no consensus exists regarding an acceptable or sufficient degree of validity for a particular purpose, readers can draw their own conclusions whether AREM's validity, as measured by the field study, is sufficient for their use. As described earlier, testing showed that AREM correctly predicted species presence or absence 88% of the time (range 40–100%). Differences in means and distributions of predicted and observed richness values were not significant. Predicted and observed species composition also was similar according to the SMC.
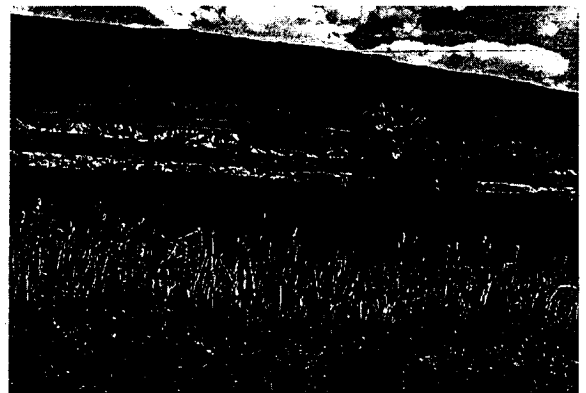
Understanding sources of correct and incorrect predictions is important for improving models, but is seldom simple or achievable without an enormous field effort (Marcot 1986, Block et al. 1994). Omission errors can represent only casual use of a habitat by a species that is merely passing through or can suggest that models defined the habitat of a species too narrowly. Commission errors can be due to un-

dersampling, to temporary species absence from an area due to ephemeral biotic or climatic factors, or to models that specify habitat preferences too broadly. Frequency of null matches, like frequency of commission errors, can be inflated by undersampling or by temporary absence of a large proportion of species from an area due to ephemeral biotic or climatic factors, especially if models specify habitat preferences too narrowly. The breeding season usually represents the best case for validating a habitat evaluation method because birds are most sedentary at that time. Testing of AREM at other seasons, using slightly different protocols, would encompass models of some other species and is needed.

Successful application of any method also hinges on the method's repeatability. Potential users of AREM (field biologists from several state and federal agencies) were involved during an earlier phase of this project in a structured testing of AREM's repeatability, with generally favorable results (Adamus 1993b). Results were used to refine the field form before its use in this study.

Management implications

AREM is intended to be intermediate in complexity between the simple WHR models and the more detailed procedures such as HEP. AREM can replace or be used as an adjunct to these other evaluation tools. Evaluations generally can be performed in <30 minutes/site, and users need not identify birds because predictions are based solely on habitat characteristics. AREM is intended to incrementally improve how particular technical data are currently used in wetland decisions. It is not intended to dictate decisions, imply perfect scientific knowledge, or provide all the answers. AREM is currently applicable only to lowland riparian and wetland areas of the Colorado Plateau, but may be adaptable to other regions, taxa, and habitat types.

This study demonstrated that, even in regions with relatively few published data on wetland bird habitat preferences, methods can be developed that rapidly and accurately predict species composition and richness of breeding bird communities explicitly. This is important because considerable time and expense are often devoted to conducting multivariate studies to identify key indicators before predictive models are even proposed. Although complex empirical and simulation approaches are still vital for addressing other questions (e.g., species response to impacts), results from the current study suggest that if the evaluation objective is simply to predict existing avian species composition and richness with general accuracy, an approach such as that described herein may be useful. Models such as AREM might be used to assist and document resource decisions in a variety of the following ways.

Mitigation calculations

For mitigation planning, resource agencies sometimes cover-type lands that will be altered or restored. This consists of measuring categories of habitat before a project and estimating shifts in area among categories as a result of the project. Areas of each cover-type category that exist both before and after the project are adjusted by coefficients determined from HEP. Where wetland and riparian cover types are the habitats that are expected to change, AREM might be used in lieu of (or in addition to) HEP to calculate the habitat suitability coefficients of affected or restored areas.

Detecting impaired wetland quality.

Where wetlands are officially considered by agencies to be waters of a state or where they exist within certain public trust lands (e.g., National Wildlife Refuges), there can be a legal need to determine the degree wetland quality has been impaired. AREM models may assist in identifying impairment by defining which species should be present in a wetland having a particular habitat. If properly designed surveys fail to detect the predicted species, it raises a possibility that factors unmeasured by AREM are discouraging wetland use by birds. Some caution is necessary because species absence could relate to weather conditions, demographic factors (e.g., suitable habitats having reduced populations of migrants because of habitat alteration on wintering areas), or weaknesses in some species models that comprise AREM. Nonetheless, AREM could be useful as an initial screening tool to help decide whether more effort should be committed to verify a problem.

Targeting habitat enhancements

Wildlife managers often seek to alter wetland conditions to improve habitat for particular species.

AREM can help identify habitat features whose enhancement will support the largest variety of species overall or a particular species. A low score for a species suggests that it might benefit from management; examination of its model sometimes indicates the conditions that could be changed or created.

Wildlife-based classification of wetland habitats

Wetland types are commonly defined by their vegetative communities. Wildlife communities or individual species also can be a useful primary or secondary feature in classifying wetlands for scientific or administrative purposes. AREM can assist such classifications by predicting bird species associated with vegetation and other environmental factors. Bird community classes of wetlands could be identified by applying AREM to a statistical sample of wetlands in a region.

Protecting biodiversity

Resource agencies and conservation groups sometimes have opportunities to purchase or trade properties to enhance regional biodiversity. When biological survey data from the subject properties are lacking, AREM can be applied (at any season) to the properties to predict their avian richness, which is often the largest terrestrial component of a region's vertebrate biodiversity. Richness estimates can be pooled from multiple wetlands to determine combinations of wetlands likely to support the greatest species richness. As such, AREM may provide a complementary, local refinement of the gap analysis approach for ecosystem management and biodiversity planning (Scott et al. 1993).

Literature cited

Adamus, P. R. 1993*a*. User's manual: Avian richness evaluation method (AREM) for lowland wetlands of the Colorado Plateau.

U.S. Environ. Prot. Agency EPA/600/R-93/240 (Natl-. Tech. Inf. Serv. # PB94-501822GEI). 72pp.

ADAMUS, P. R. 1993*b*. Irrigated wetlands of the Colorado Plateau: information synthesis and habitat evaluation method. U.S. Environ. Prot. Agency EPA/600/R-93/071 (Natl. Tech. Inf. Serv. # PB93-186260). 91pp.

BLOCK, W. M., M. L. MORRISON, J. VERNER, AND P. N. MANLEY. 1994. Assessing wildlife-habitat-relationships models: a case study with California oak woodlands. Wildl. Soc. Bull. 22:549-561.

DEXTER, C., AND R. LAVAD. 1992. Bird check list for the Grand Valley and surrounding high country of Mesa County, Colorado. Grand Valley Audubon Society, Grand Junction, Colo. 2pp.

DIGBY, P. G. N., AND R. A. KEMPTON. 1987. Multivariate analysis of ecological communities. Chapman and Hall, London, U.K. 206pp.

GARRISON, B. A. 1993. Validation studies of the California wildlife habitat relationships system. Fish and Wildl. Inf. Exchange Newsl. 2(2):2-5.

GLAHN, J. F. 1974. Study of breeding rails with recorded calls in north-central Colorado. Wilson Bull. 86:206-214.

GOLET, F. C. 1973. Classification and evaluation of freshwater wetlands as wildlife habitat in the glaciated Northeast. Ph.D. Diss., Univ. Massachusetts, Amherst. 179pp.

HURLEY, J. F. 1986. Summary: development, testing, and application of wildlife-habitat models—the researcher's viewpoint. Pages 151-154 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. Univ. Wisconsin Press, Madison.

KINGERY, H. E., EDITOR. 1988. Colorado bird distribution latilong study. Revised ed. Colo. Div. of Wildl., Denver. 81pp.

MARCOT, B. G. 1986. Summary: biometric approaches to modeling—the manager's viewpoint. Pages 203-204 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. Univ. Wisconsin Press, Madison.

MORRISON, M. L., B. G. MARCOT, AND R. W. MANNAN. 1992. Wildlife-habitat relationships. Univ. Wisconsin Press, Madison. 343pp.

PRESTON, F. W. 1948. The commonness, and rarity, of species. Ecology 29:254-283.

RECTOR, C. D., E. W. MUSTARD, AND J. T. WINDELL. 1979. Lower Gunnison River Basin wetland inventory and evaluation. U.S. Dep. Agric. Soil Conserv. Serv., Denver, Colo. 90pp.

RICE, J. C., R. D. OHMART, AND B. W. ANDERSON. 1986. Limits in a data-rich model: modeling experience with habitat management on the Colorado River. Pages 79-86 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. Univ. Wisconsin Press, Madison.

RICKLEFS, R. E., AND D. SCHLUTER, EDITORS. 1993. Species diversity in ecological communities. Univ. Chicago Press, Chicago, Ill. 414pp.

SCOTT, J. M., F. DAVIS, B. CSUTI, R. NOSS, B. BUTTERFIELD, C. GROVES, H. ANDERSON, S. CAICCO, F. D'ERCHIA, T. C. EDWARDS, JR., J. ULLIMAN, AND R. G. WRIGHT. 1993. Gap analysis: a geographic approach to protection of biological diversity. Wildl. Monogr. 123. 41pp.

SOKAL, R. R., AND C. D. MICHENER. 1958. A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 38:1,409-1,438.

U.S. BUREAU OF RECLAMATION. 1991. Wetlands inventory and evaluation, Uncompahgre Project System optimization/joint use. U.S. Bur. of Reclamation, Upper Colorado Reg., Grand Junction, Colo. 32pp.

U.S. FISH AND WILDLIFE SERVICE. 1980. Habitat evaluation procedures (HEP) manual (102ESM). U.S. Fish and Wildl. Serv., Washington, D.C. 144pp.

VAN HORNE, B., AND J. A. WIENS. 1991. Forest bird habitat suitability models and the development of general habitat models. Fish and Wildlife Research Rep. 8. U.S. Fish and Wildl. Serv., Washington, D.C. 31pp.

Paul R. Adamus is Project Scientist for ManTech Environmental Research Services Corporation in Corvallis, Oregon. He received his B.S. in Wildlife Science from the University of Maine and an M.S. in Biology from the University of Utah. Paul contributes to the U.S. Environmental Protection Agency's Wetlands Research Program and specializes in wetland functional assessment. He also coordinates Oregon's statewide Breeding Bird Atlas Project.

Associate Editor: Leopold.